

Comparing two teacher development programs for innovating reading comprehension instruction with regard to teachers' experiences and student outcomes[☆]

Hilde Van Keer*, Jean Pierre Verhaeghe

Department of Education, Ghent University, Ghent, Belgium

Abstract

This study compared a year-round intensive coaching of teachers (35 h) with a restricted in-service course (13 h), designed on the basis of research-based components of effective professional development of teachers. Both courses were developed to innovate reading comprehension instruction in elementary schools. Fourteen second- and 16 fifth-grade teachers participated. Results indicated that both conditions were equally effective in changing students' reading comprehension, reading fluency, reading strategy use, and self-efficacy perceptions. Teachers' experiences with the professional development course were also comparable in both conditions. A significant point of difference indicated; however, that teachers attending the restricted in-service course experienced more workload with regard to settling in the innovations.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: In-service professional development of teachers; Reading comprehension instruction; Elementary school teachers

1. Introduction

Learning to read is one of the most crucial learning processes children are involved in elemen-

tary school. However, especially with regard to the ultimate goal of reading comprehension many children seem to have rather persistent difficulties. Reading comprehension can be defined as constructing a mental representation of textual information and its interpretation (Van Den Broek & Kremer, 2000) or, in other words, as extracting meaning from written words, sentences, and texts (Aarnoutse & Van Leeuwe, 2000). Notwithstanding the apparent simplicity of this definition, it is a complex process. Cognitively based views of reading comprehension emphasize that proficient readers do much more than just word-, phrase-, or

[☆]The present study is part of a larger long-term study investigating the effects of a program innovating reading comprehension instruction in Flanders (Belgium) on the cognitive, social, and emotional development of elementary school children. The study was supported by a Research Assistantship Grant of the Fund for Scientific Research—Flanders.

*Corresponding author. Tel.: 32 9 264 86 62.

E-mail address: Hilde.VanKeer@UGent.be (H. Van Keer).

sentence-level processing; the mastery and use of both metacognitive and cognitive strategies that facilitate text comprehension distinguishes poor and skilled readers (Baker & Brown, 1984; Dole, Duffy, Roehler, & Pearson, 1991). Notwithstanding the importance of reading strategies as tools for enhancing text comprehension, there is no reason to assume that all elementary students spontaneously discover them and appeal to strategic processes when confronting texts that are challenging to comprehend (Hartman, 2001; Pressley & Allington, 1999). Nevertheless, a substantial body of research results reveals that even when children do not use effective comprehension strategies on their own, explicit reading strategies instruction is a feasible tool for teaching students to acquire and apply them successfully (e.g., Brand-Gruwel, Aarnoutse, & Van Den Bos, 1998; De Corte, Verschaffel, & Van De Ven, 2001; Dole et al., 1991; Klingner & Vaughn, 1996; Pressley, 2000; Van Keer, 2004; Van Keer & Verhaeghe, 2003). In addition to the often cited importance of explicit instruction in reading strategies, learning to read appears to be a social process. Research convincingly establishes that interaction among peers about texts promotes comprehension and the application of self-regulation reading strategies (e.g., Almasi, 1996; Brown, Pressley, Van Meter, & Schuder, 1996; Dole et al., 1991; Fuchs & Fuchs, 2000; Fuchs, Fuchs, Mathes, & Simmons, 1997; Klingner, Vaughn, & Schumm, 1998; Mathes & Fuchs, 1994; Palincsar & Brown, 1984; Pressley et al., 1992; Rosenshine & Meister, 1994; Simmons, Fuchs, Fuchs, Mathes, & Hodge, 1995; Van Keer, 2004; Van Keer & Verhaeghe, 2003). Despite these research results documenting that peer-led interaction about texts and explicit instruction in the strategic aspects of processing and comprehending text are more beneficial in promoting students' strategic reading and reading comprehension than traditional comprehension instruction, the successful practices emanated from the aforementioned intervention studies, unfortunately were not translated into educational reality. The practice of teaching reading nowadays is still very traditional, mainly characterized by teacher-led comprehension "testing" or questioning students about the content of a

text after reading it and hardly any overt and continuous strategy instruction (Aarnoutse, 1995; Aarnoutse & Weterings, 1995; Dole, 2000; Paris & Oka, 1986; Pressley, Wharton-McDonald, Hampson, & Echevarria, 1998) or student-centered discussion (Alvermann, 2000). Taken into account this marked gap between empirical research and instructional practice, a major issue of concern is the development and implementation of effective ways to prepare teachers to tune their teaching to recent research findings. Therefore, researchers should accept the challenge to inquire into effective strategies for disseminating research-based practices, for questions remain concerning how to conceptualize teacher learning and, correspondingly, about how to construct professional development so as to foster meaningful change in educational practice (Butler, Lauscher, Jarvis-Selinger, & Beckingham, 2004).

2. Study's purpose and hypothesis

The aim of the present study was to compare the efficacy of two variants of professional development of teachers as strategies to update teachers' knowledge on recent research developments in the field of reading comprehension and to translate this knowledge in classroom practice with the intention of making reading comprehension instruction in elementary schools more effective. Both variants were composed on the basis of research-based effective components of teacher development, but distinguished themselves by the amount of support lent to the teachers. Former research already confirmed the positive impact of an intensive year-round counselling, coaching, and on-the-job-development of elementary school teachers on their students' reading comprehension, reading fluency, and self-efficacy perceptions towards reading (Van Keer, 2004; Van Keer & Verhaeghe, 2003, 2004). The extensive assistance provided to the teachers was; however, considered as a restriction of the study, for it is unlikely that regular in-service authorities, responsible for the professional development of teachers, would be able to offer commensurate levels of support. In this respect, the present study intends to meet the

plea for developing a concise teacher development course, of which the effects on teachers and student outcomes would be comparable with those of an intensive year-round on-the-job-development of teachers. More specifically, a year-round elaborated coaching of teachers (35 h) was equated to a more restricted 13-h in-service course. Since both the intensive and restricted in-service course were composed on the basis of research-based effective components of teacher development, it was hypothesized that both teacher development strategies would be equally effective, so teachers' experiences as well as student outcomes would be comparable in both development courses. Therefore, we expected no significant differences between the intensive coaching condition and the restricted in-service course condition on the measures of teachers' reported experiences, nor on students' pre- to post- and retention test evolution with regard to reading comprehension, reading fluency, use of reading strategies, and self-efficacy perceptions towards reading.

3. Description of the strategies for professional development of teachers

This section outlines the two variants of teacher development developed to assist elementary schools in making reading comprehension instruction more effective. At first the most important aspects of the procedures and principles of the innovative approach treated in the development will be highlighted briefly. For more detailed information, we invite the reader to consult former publications (Van Keer, 2002, 2004; Van Keer & Verhaeghe, 2003). Subsequently, the organization of both teacher development variants will be presented.

3.1. The innovative approach towards reading comprehension instruction

The aforementioned research-based practices of explicit reading strategies instruction and regular student-led interactions about texts were blended in the innovative approach dealt with in the teacher development course. Weekly organized

peer tutoring activities were chosen as the instructional technique to create opportunities to stimulate student-centered talk about the interpretation of texts and to practice the application of the explicitly taught reading strategies. The discussed innovative approach towards reading comprehension instruction was identical in both variants of teacher development. More specifically, the intended instructional innovations were characterized by three components: explicit reading strategies instruction, a sound tutor preparation, and weekly organized cross-age or reciprocal same-age peer tutoring sessions. With regard to the first component, the innovation focused on the acquisition and mastery of six reading strategies by means of modeling strategic reasoning and explicit teacher explanations of why, where, and when to use them. In addition, a gradual transfer from external regulation by the teacher to self-regulation of strategy use by the students was taken into account. Since there is evidence that peer tutoring is less effective when no attention is paid to a sound preparation of the tutors (Bentz & Fuchs, 1996; Fuchs, Fuchs, Bentz, Phillips, & Hamlett, 1994; Fuchs et al., 1997), a series of preparatory lessons and student materials focusing on the acquisition of social, communicative, and procedural skills was developed and included as the second element of the innovative approach. Finally, the innovations were typified by weekly peer tutoring activities to systematize the use of the explicitly taught reading strategies by means of structured assignments cards. Peer tutoring activities were organized class-wide, so all students in a class were paired and worked simultaneously.

3.2. Intensive year-round coaching

Teachers in the intensive year-round coaching condition were provided with an elaborated manual giving step by step instructions on how to conduct the instructional innovations. The manual included an extensive general description of the rationale, aims, and the organization of the innovations; lesson scenarios describing the objectives, the necessary materials, the preferable instructional techniques, and the successive phases of each lesson; and supplementary students

materials, such as strategy assignment cards and reading texts. The manual was developed by the first author of the article, based on extensive review of related empirical research (e.g., Bentz & Fuchs, 1996; Brown et al., 1996; Fuchs et al., 1997; Fukkink, Van der Linden, Vosse, & Vaessen, 1997; Klingner & Vaughn, 1996; Klingner, Vaghn & Schumm, 1998; Palincsar & Brown, 1984; Pressley et al., 1992). Teachers were not required to develop additional materials.

In addition to the manual, teachers were provided with an intensive year-round counselling, coaching, and on-the-job-development by the first author of the article. The coaching activities took place at the local schools and were launched with a 3-h introductory session prior to the start of the implementation of the innovations. Since teachers need evidence that the innovative approaches do actually work (Butler et al., 2004; Stein, Schwan Smith, & Silver, 1999), this meeting focused on developing understanding of the characteristics of reading instruction that successfully improve students' comprehension of texts and on the general organization of the three innovative instructional components. A short video film made the innovations more concrete and outlined how to introduce them in the regular instructional practice. As recommended by Loucks-Horsley, Hewson, Love, and Stiles (1998) teachers were in this respect helped to translate their knowledge of the innovative practices into practice. Clarke and Hollingsworth (2000, 2002) also recognize the need to contextualize teaching and teacher development by employing cases, and more specifically video cases, as a means to situate the professional development of teachers in realistic contexts. In addition, room was left for questions and discussion. After the first introduction, the lesson scenarios and additional student materials with regard to the tutor preparation and the explicit reading strategies instruction were fully discussed during six meetings. These meetings more specifically comprised an observation of a preparatory or explicit strategy instruction lesson, a detailed discussion of the attended and other already completed lessons, and a preparation of the coming lessons. Again a video film was employed to concretize the interpretation of the lesson

scenarios. Prior to the first peer tutoring activity, an additional meeting was organized to discuss the composition of the reading dyads, the practical organization of the tutoring sessions, and the necessary arrangements for a smooth course. Moreover, precisely selected extracts from the video film were used to stimulate discussion and debate about the importance of the teachers' role in supervising and coaching the reading dyads. Finally, during the implementation of the peer tutoring activities, monthly in-class observations took place, followed by discussions to exchange experiences and ideas, and to overcome practical or implementation difficulties. In general, teachers were coached during approximately 35 h, spread out over the entire school year.

3.3. Restricted 13-hours in-service course

Teachers in the restricted 13-h in-service course condition were provided with an identical manual and the same supplementary student materials as the teachers in the coaching condition. However, the additional assistance and support offered to the teachers was less extensive, geared to the levels of support that regular in-service authorities responsible for the professional development of teachers are able to offer. Since the aim of the present study was to develop a less intensive teacher development with effects comparable to the year-round teacher coaching, the program was based on essential elements in in-service courses increasing the likelihood of teacher implementation.

The predominant strategy for professional development for teachers continues to be the use of short in-service workshops (McCutchen & Berninger, 1999; Veenman, Van Tulder, & Voeten, 1994), mainly characterized by a top-down approach to disseminate knowledge, in which teachers are provided with information and resources that they are expected to translate into action (Gersten, Vaughn, Deshler, & Schiller, 1997; Little, 1993). With this approach, teachers are not seen as active participants in their own professional growth (Sandholtz, 2002). Notwithstanding their popularity; however, there is considerable evidence that these one-shot teacher

development strategies foster only surface level or shallow post-course implementation of the intended changes in the classroom as opposed to deep rooted changes in practice (Ball & Cohen, 1996; Bradley, Conner, & Southworth, 1994; Englert & Tarrant, 1995; Fullan & Stiegelbauer, 1991; Hawley & Valli, 1999; Joyce & Showers, 1980, 1982; Leach & Ingram, 1989; Wade, 1985). Teachers themselves describe these one-shot workshops as boring and irrelevant, and they tend to forget 90% of what they learn (Miller, 1998). Follow-up observations rarely show that teachers use the new knowledge in their daily teaching repertoires, even when those innovations are effective (Gersten et al., 1997). Moreover, measures of change in student outcomes have also rarely been reported. Therefore, it can be concluded that in-service development is all too often ineffective and that the presentation of information and discussion are necessary, but not sufficient conditions to bring about change. An important limitation of this traditional model of professional development is the passive role imposed upon teachers and the separation from teachers' daily practice. Teachers tend to experience professional growth as a separate and distinct event, suggesting that teachers' learning can be segmented from their regular work. Rather than drawing upon teachers' classroom practice, professional development is frequently disconnected from it both in substance and setting (Bell, 1991; Burbank & Kauchak, 2003; Sandholtz, 2002). Research, however, identified a combination of in-service development characteristics that do increase the probability of post-course teacher implementation, for example, focused presentation of information and discussion about new strategies, theories or ideas to be applied, in such a format that they become acceptable and usable for the teachers; clear identification of target behaviors to be changed; demonstration and practice of relevant behaviors to be implemented; structured opportunities for reflection on teaching and learning; and post-course, in-class follow-up procedures, such as supervised practice and consultation, and performance feedback to the teachers (Djalil & Anderson, 1989; Harchik, Sherman, Sheldon, & Strouse, 1992; Loucks-Horsley et al.,

1998; Norman, Sprinthall, & Ties-Sprinthall, 1996; Showers, 1990; Veenman et al., 1994; Wheldall, Merrett, & Borg, 1985).

Taken into account the identified effective components of professional development of teachers, the restricted in-service course developed in the framework of the present study comprised an informative part, as well as in-class follow-up meetings with performance feedback to the teachers. Both components were conducted by the first author of the article and were organized at the local schools. More specifically, teachers were offered three 3-h preparatory local school meetings directed at the presentation of information on the elements in the innovative approach and at the identification and demonstration of the instructional behaviors to be implemented. In these meetings teachers were required to participate actively in the discussions and to formulate feasible plans for implementing the innovative approach in their own reading lessons. The active participation of the teachers in the meeting was emphasized since collaborative models of teacher development stress the importance of nurturing learning communities within which teachers try out new ideas in the classrooms, reflect on outcomes, and co-construct knowledge about teaching and learning in the context of authentic activity (Butler et al., 2004; Englert & Tarrant, 1995; Palincsar, Magnusson, Marano, Ford, & Brown, 1998; Perry, Walton, & Calder, 1999). In addition, we followed the teachers back into their classes during two follow-up booster sessions, characterized by observation and consultation in the teachers' classrooms and a discussion afterwards. The first preparatory school meeting ran completely parallel with the 3-h introductory session in the intensive coaching condition. The session more particularly focused on the presentation of information and discussion concerning the importance of explicit reading comprehension instruction, the organization of peer tutoring activities as a vehicle to encourage interaction about texts, and the significance of a sound tutor preparation. The underlying theoretical rationale and the general organization of the three innovative components were clarified and a short video film demonstrated the intended innovations. The

second and third preparatory meeting dealt with the identification of the instructional behaviors to be changed by means of a discussion of the interpretation of the lesson scenarios and additional student materials, respectively, with regard to the tutor preparation and the explicit reading strategies instruction. Again, the presentation of the intended instructional innovations was interspersed with demonstrations by means of video fragments. The third local meeting was completed with a discussion of the composition of the reading dyads, the practical organization of the peer tutoring activities, the necessary managerial teacher behaviors and arrangements for a smooth course, and the importance of the teachers' role in supervising and coaching the teams. Precisely selected video film extracts of reading dyads in action were used to stimulate discussion and debate and to encourage teachers to practice and express how to react and intervene to coach the teams. Finally, the restricted in-service teacher development course was completed with two 2-h in-class follow-up sessions. These sessions started with an observation of a peer tutoring session. Thereafter, a discussion with the teachers was organized to go into the attended lesson, to give performance feedback based on the observational data, to exchange experiences and ideas, and to overcome practical or implementation difficulties. These structured opportunities for reflection are central, since it is difficult to make meaningful shifts in practice without stepping away from immediate demands or having time to reflect on teaching (Butler et al., 2004).

4. Method

4.1. Participants

The present study is part of a larger long-term study that investigated the effects of a program innovating reading comprehension instruction in Flanders (Belgium). Within this framework, regular elementary school teachers were recruited 2 year before by an informative article in two widespread journals accompanied with an invitation to subscribe to a professional development

program for teachers on innovative reading comprehension instruction. More specifically, we focused on second- and fifth-grade teachers. Approximately 100 teachers volunteered to be engaged. Based on their responses to a comprehensive questionnaire concerning teachers' teaching orientation (Maes, Vandenberghe, Ghesquière, & Luyckx, 1997), instructional techniques (Reezigt, 1993), and learning preference (Owens & Barnes, 1992), student-oriented teachers with some experience in cooperative and interactive instructional techniques were selected for participation. Further selection was based on the geographical distribution of schools throughout the whole of Flanders and on the possibility to match teachers and classes in the different research conditions with regard to teachers' teaching experience, beliefs, and instructional practice, class size, and students' age, gender distribution, and dominating mother tongue. More specifically, 14 second- and 16 fifth-grade teachers from 16 different schools were selected to participate. They were all white and had Dutch as their mother tongue. Second- and fifth-grade teachers had on average 13.25 (SD = 4.37) and 22.29 (SD = 9.09) years of teaching experience. Only one second-grade and 5 fifth-grade teachers were men. At the time of the present research, the teachers participated for the second year in the larger long-term study. In the previous school year the teachers in the intensive coaching condition had received a similar elaborated coaching aimed at implementing the explicit reading strategies instruction, as well as the peer tutoring activities. The teachers attending the restricted professional development course had been involved in a less elaborated in-service program directed at introducing explicit reading strategies instruction and practicing strategic reading in teacher-led whole-class activities.

All students of the participating 30 teachers were involved in the study, yielding in total 272 second and 342 fifth graders. Informed consent was obtained of all students' parents via the schools principals. With the exception of one inner-city school with mainly a low SES and ethnic minority population, the schools' population was comprised chiefly of white, Flemish students from middle-class families. Except for

Table 1
Number of participating classes and pupils

Condition	Grade			
	2nd		5th	
	Classes	Students	Classes	Students
Intensive year-round coaching	7	139	7	146
Restricted in-service training	7	133	9	196
Total	14	272	16	342

one fifth-grade class including only boys, there was approximately an even division of gender: on average 50% ($SD = 19.19$) of the second- and 56% ($SD = 20.61$) of the fifth-grade students were boys. The majority of the students (237 in second and 325 in fifth grade) were native Dutch speakers. Classes are to be considered as academically heterogeneous. Class size ranged from 10 to 26 with an average of 19 ($SD = 3.99$) in second grade and from 9 to 34 in fifth grade with on average 21 ($SD = 6.88$) students per class, which is representative for the Flemish situation. Exceptionally small and large classes are the exception rather than the rule and were equally distributed across treatments.

From the start of the long-term study onwards, participating teachers were randomly assigned to the intensive year-round coaching or restricted in-service development condition. Table 1 presents the number of second- and fifth-grade classes and students per condition. Univariate analysis of variance revealed no significant differences between both conditions with regard to second- and fifth-grade teachers' teaching experience ($F = 0.47$; $df = 1$; $p = 0.508$; $F = 1.84$; $df = 1$; $p = 0.200$), class size ($F = 0.10$; $df = 1$; $p = 0.757$; $F = 0.10$; $df = 1$; $p = 0.758$), gender distribution ($F = 0.63$; $df = 1$; $p = 0.442$; $F = 1.13$; $df = 1$; $p = 0.306$), and percentage of non-native speakers ($F = 0.37$; $df = 1$; $p = 0.552$; $F = 0.22$; $df = 1$; $p = 0.646$).

4.2. Instruments

To compare the effectiveness of both conditions for professional development of teachers, teachers' experiences as well as student outcomes were

assessed. Teachers' experiences were measured by means of a questionnaire concerning satisfaction with the development course, perception of workload associated with implementing the innovative instructional approach, and appraisal of student progress by the innovations. More specifically, the questionnaire was developed within the framework of the present study and incorporated 45 representative multiple-choice questions, all with a five-point Likert-type response format. With regard to student outcomes, standardized reading comprehension and fluency tests, as well as questionnaires assessing the use of reading strategies and pre-occupation with attributions and self-efficacy perceptions towards reading were administered at three different measurement occasions. The present section first outlines the different parts of the teacher questionnaire. Further, the measurement instruments with regard to student outcomes will be described briefly.

Questionnaire concerning teachers' satisfaction with the development course: Four items of the questionnaire assessed teachers' satisfaction with the attended development course. More specifically, they were asked to rate their overall contentment with taking part in the course, as well as their satisfaction with the information and counselling about the innovative instructional program, the provided manual, and the additional student materials. The response format ranged from "strongly satisfied" (1) to "strongly dissatisfied" (5).

Questionnaire concerning teachers' perception of workload: Eleven statements examined the workload teachers experienced as a result of implementing the innovative instructional approach in their regular teaching practice. The correlations between the statements showed significant overlap between certain groups of items. Therefore, responses were subjected to principal components analysis using varimax rotation. This analysis revealed three underlying factors. Five items loaded high on factor one, conceptually representing teachers' experience of workload with regard to the implementation and organization of peer tutoring activities. Three items loaded on the second factor, reflecting the workload in reference to settling in the innovative approach and preparing the

students for practicing the application of reading strategies in peer tutoring dyads. For both these factors the response format of the items ranged from “experienced as a strong load” (1) to “absolutely not experienced as a load” (5). Finally, three items loaded on factor three, representing teachers’ impression of increased opportunities to differentiate according to pace or content for reading dyads of different abilities as a result of organizing peer tutoring activities. In this case, the response format of the statements ranged from “opportunities to differentiate strongly experienced” (1) to “absolutely no opportunities to differentiate experienced” (5). Factor one accounted for almost 35% of the variance and factor two explained 32% of the variance. The third factor accounted for an additional 14% of the variance. In sum the three factors explained almost 82% of the variance observed in the 11 survey items. To verify the reliability of the three subscales Cronbach’s α -coefficients were computed, which revealed reliable measures for the first ($\alpha = 0.90$; $n = 21$), the second ($\alpha = 0.91$; $n = 19$), as well as the third factor ($\alpha = 0.86$; $n = 21$). Subscale scores were computed by averaging out the responses to the items within the factors. These scores were used in the subsequent analyses aimed at comparing the experiences of teachers from both conditions.

Questionnaire concerning teachers’ appraisal of student progress: Thirty statements focused on teachers’ appraisal of student progress as a result of implementing the innovative instructional approach. The response format with regard to these statements ranged from “students progressed very strongly” (1) to “no progress observed” (5). More specifically, this part of the questionnaire was divided into different components, respectively, assessing students’ progress in the field of reading comprehension and fluency, social and interaction skills, approach to assignments, and attitude towards reading. In four items teachers were asked to rate students’ overall progress in the field of reading comprehension, as well as their advancement in the knowledge, the application, and the spontaneous use of reading strategies. In a fifth statement they evaluated students’ reading fluency improvement. These five items were treated as separate dependent variables.

Fourteen items examined teachers’ evaluation of student progress in the field of social and interaction skills. Principal components analysis using varimax rotation was applied to reduce these statements to a more feasible number of underlying subscales. Three factors arose, respectively, accounting for almost 48%, 25%, and 9% of the variance in the original set of items. Five items loaded on factor one, representing advancement in general social and interaction skills; five items loaded on the second factor, representing evolution in social relationships and friendships. Four items loaded on the last factor, reflecting progress in the application of social and interaction skills during the peer tutoring activities. The three factors were examined for reliability. All showed acceptable internal consistency. Cronbach’s α -coefficients were, respectively, 0.84 ($n = 19$), 0.88 ($n = 18$), and 0.85 ($n = 17$). Subscale scores were computed by averaging out the responses to the respective items and used in the subsequent analyses.

Eight items focused on teachers’ evaluation of students’ approach to assignments. More specifically, changes in responsibility, self-confidence, and autonomy were rated. Again, responses were subjected to principal components analysis using varimax rotation. This analysis revealed two underlying factors, reflecting students’ approach to tasks in general and to peer tutoring assignments more specifically. Three items loaded high on factor one, five items loaded on the second factor. Factor one accounted for 51% of the variance and factor two explained 23% of the variance observed in the originally survey items. To verify the reliability of the three subscales, Cronbach’s α -coefficients were computed, which revealed reliable measures for the first ($\alpha = 0.94$; $n = 19$) and second factor ($\alpha = 0.79$; $n = 16$). Factor scores were computed by averaging out the responses.

Finally, three items dealt with teachers’ assessment of changes in students’ attitude towards reading. Since the Cronbach’s α -coefficient of the combination of these items was highly reliable ($\alpha = 0.87$; $n = 22$), the three response scores were averaged out and used in subsequent analyses.

Standardized reading comprehension tests: Students' reading comprehension achievement was measured using widely used, well-established, and valid Dutch standardized test batteries, addressing aspects that are covered by the reading strategies taught in the innovative instructional reading comprehension program. For the pre-test in second grade a test from the battery "Lezen met begrip 1" (Reading with comprehension) (Verhoeven, 1993) was utilized. For the post- and retention test, two different tests from the battery "Toetsen Begrijpend Lezen" (Reading comprehension tests) (Staphorsius & Krom, 1996) were administered. The students read short expository and narrative texts and then answered multiple-choice questions with four alternatives. Scores were determined by summing up the correct answers. Cronbach's α -coefficients yielded high reliability scores of 0.90 ($n = 270$) for the pre-test (30 items), 0.80 ($n = 264$) for the post-test (25 items), and 0.81 ($n = 238$) for the retention test (25 items).

In fifth grade reading comprehension achievement was measured with the standardized test battery "Toetsen Begrijpend Lezen" (Reading comprehension tests) (Staphorsius & Krom, 1996). The tests consisted of three modules of 25 multiple-choice questions with four alternatives each. All students took the first module of the test. Afterwards they completed the second more easy or third more difficult module, depending on their first results. Scores were determined by summing up the correct answers and transposing the sum score into an IRT-modeled global achievement score. To verify the reliability of the three modules of the pre-, post-, and retention test, Cronbach's α -coefficients were computed. Table 2 indicates that all reading comprehension measures are acceptably reliable. Only the most difficult module of the retention test shows a relatively low internal consistency level, which probably can be attributed to the difficulty of the test and the small number of children that completed this module.

Decoding fluency test: The Dutch standardized "Eén minuut test" (One minute test) (Brus, 1969) was used at each measurement occasion to assess second graders' reading fluency. More specifically, during exactly 1 minute individual students read words out of a list of 100 unrelated words with a

Table 2

Cronbach's α -coefficients of the Fifth Grade Reading Comprehension Tests

Test module	Measurement occasion ^a		
	Pre-test	Post-test	Retention test
1	0.82 ($n = 340$)	0.75 ($n = 316$)	0.76 ($n = 280$)
2	0.73 ($n = 122$)	0.83 ($n = 177$)	0.76 ($n = 256$)
3	0.71 ($n = 205$)	0.74 ($n = 130$)	0.52 ($n = 25$)

^aAt each measurement occasion a different test, increasing in level of difficulty, was used.

gradually increasing level of difficulty. The final score was determined by counting the number of words read correctly. The "Eén minuut test" is one of the most widely used, valid, and reliable decoding fluency test in the Dutch linguistic area.

Questionnaire on the use of reading comprehension strategies: To estimate to what extent second- and fifth graders use reading strategies, a questionnaire was developed and administered at each measurement occasion to both second and fifth graders. More specifically, it concerns a list of 20 statements with respect to the use of strategies before, while, and after reading. Students were asked to report how often the statements applied to their own reading behavior by ticking one of the boxes ("never", "seldom", "sometimes", "very often or always"). Examples of the statements are: "Before I start reading, I consider what I already know about the subject of the text.", "While reading, I look for the meaning of difficult words", "While reading, I try to figure out what the story is about." Cronbach's α -coefficients are presented in Table 3 and reveal that the questionnaire is acceptably reliable for both second and fifth graders.

Questionnaire on self-efficacy perceptions and related causal attributions: A questionnaire was developed to measure how often students are pre-occupied with positive or negative thoughts with regard to their reading ability or related causal attributions. Children were asked to report how often such thoughts crossed their mind either before, while, or after reading by ticking off "never", "seldom", "sometimes", or "very often or always", following statements as: "I think: I

Table 3

Cronbach's α -coefficients of the Questionnaire concerning Preoccupation with Attributions and Self-efficacy Perceptions

Questionnaire scale	Measurement occasion					
	Pre-test		Post-test		Retention test	
	2nd grade	5th grade	2nd grade	5th grade	2nd grade	5th grade
Use of reading comprehension strategies	0.67 ($n = 227$)	0.81 ($n = 280$)	0.77 ($n = 230$)	0.87 ($n = 269$)	0.74 ($n = 193$)	0.87 ($n = 248$)
Success attributions and positive self-efficacy perceptions	0.60 ($n = 247$)	0.68 ($n = 320$)	0.69 ($n = 258$)	0.67 ($n = 302$)	0.67 ($n = 225$)	0.71 ($n = 266$)
Failure attributions and negative self-efficacy perceptions	0.74 ($n = 235$)	0.81 ($n = 291$)	0.76 ($n = 239$)	0.86 ($n = 278$)	0.83 ($n = 213$)	0.86 ($n = 247$)

comprehended well because it was an easy text.”, “I think: I did not comprehend well because there was no one to help me.”, “I think: I am not good at reading”. Former research (Van Keer & Verhaeghe, 2003) revealed that success attributions and positive thoughts about one's reading competence on the one hand and failure attributions and negative self-efficacy perceptions with regard to reading on the other hand are very closely related. Therefore, two scales were constructed. The questionnaire was administered at each measurement occasion to both second and fifth graders. Table 3 presents the results of the internal consistency analyses, revealing that both scales are acceptably reliable.

4.3. Procedure

For both conditions teachers' development started in September, at the start of the school year. After the first introductory sessions, teachers started the implementation of the innovative instructional approach. All teaching activities involved in the innovative approach were conducted by the regular classroom teachers during time normally allocated for reading instruction, for one or two periods a week during the entire school year. Teachers completed the questionnaire assessing their experiences at the end of the school year. Student data were collected within the regular classroom context and during regularly scheduled class sessions. Student tests were administered

at three points in time: a pre-test in October (second and fifth grade) before the implementation of the innovations, a post-test in May or June (second and fifth grade) after the completion of the innovations, and a retention test in December (third and sixth grade).

5. Results

5.1. Experiences reported by the teachers

Table 4 presents the average teacher response scores on the different parts of the questionnaire and their subscales. More specifically, the averages for both teacher development strategies in both grades, as well as the overall mean are displayed. To investigate whether teachers' experiences differed significantly according to grade level (second versus fifth grade) and the professional development they attended (intensive year-round coaching versus a restricted in-service course), multivariate and univariate analysis of variances were executed.

Questionnaire concerning teachers' satisfaction with the development course: The overall mean scores in Table 4 indicate that teachers are generally strongly satisfied with attending the professional development course, especially with the information and counselling about the innovative program, and with the provided manual. The supplied student materials were somewhat less appreciated, but still reached a rather high level of

Table 4

Means and standard deviations of the teacher questionnaire results on a five-point Likert-Scale

Questionnaire variables	Grade				Total
	2nd		5th		
	Intensive coaching	Restricted course	Intensive coaching	Restricted course	
<i>Satisfaction</i>					
Overall	1.80 (0.84)	1.80 (0.45)	1.83 (0.41)	1.25 (0.50)	1.70 (0.57)
Information/counselling	1.20 (0.45)	1.40 (0.55)	1.33 (0.52)	1.25 (0.50)	1.30 (0.47)
Manual	1.40 (0.89)	1.40 (0.55)	1.33 (0.52)	1.25 (0.50)	1.35 (0.59)
Student materials	2.20 (0.45)	1.80 (0.45)	2.00 (0.63)	2.75 (0.96)	2.15 (0.67)
<i>Workload</i>					
Settle in/preparation	4.42 (0.79)	3.33 (0.67)	4.44 (0.46)	3.92 (0.74)	4.12 (0.73)
Organization tutoring	3.10 (0.50)	3.47 (0.42)	3.67 (0.73)	4.10 (0.77)	3.60 (0.69)
Chance to differentiate	2.50 (0.58)	2.67 (0.33)	3.00 (0.42)	2.17 (0.43)	2.63 (0.53)
<i>Student progress</i>					
Reading comprehension	2.75 (0.50)	3.00 (0.00)	3.20 (0.45)	2.25 (0.50)	2.81 (0.54)
Strategy knowledge	3.00 (0.00)	3.00 (0.00)	2.80 (0.45)	2.00 (0.00)	2.69 (0.48)
Strategy use	2.75 (0.50)	2.67 (0.58)	3.20 (0.45)	2.25 (0.50)	2.75 (0.58)
Spontaneous strategy use	2.75 (0.50)	3.00 (0.00)	2.80 (0.45)	2.75 (0.50)	2.81 (0.40)
Reading fluency	2.17 (0.41)	2.20 (0.45)	2.83 (0.98)	2.00 (0.00)	2.35 (0.67)
General social skills	2.47 (0.42)	2.64 (0.54)	2.28 (0.33)	2.70 (0.99)	2.49 (0.49)
Social skills in tutoring	2.25 (0.66)	2.70 (0.45)	2.10 (0.49)	1.88 (0.18)	2.30 (0.54)
Social relationships	2.80 (1.06)	2.56 (0.52)	2.48 (0.58)	2.60 (0.28)	2.59 (0.59)
General task approach	2.20 (0.45)	2.67 (0.58)	2.28 (0.44)	3.00 (0.94)	2.42 (0.55)
Tutoring tasks approach	2.32 (0.52)	2.40 (0.20)	2.20 (0.36)	2.10 (0.14)	2.26 (0.36)
Attitude towards reading	2.39 (0.44)	1.94 (0.44)	2.39 (0.49)	2.25 (0.96)	2.24 (0.57)

contentment. A 2×2 MANOVA was performed on the four satisfaction statements. Using Wilks' lambda criterion, the multivariate tests revealed no significant grade level ($F(4, 13) = 1.45$; $p = 0.272$), condition ($F(4, 13) = 1.21$; $p = 0.350$), or interaction effect ($F(4, 13) = 1.17$; $p = 0.371$). As Table 5 reveals, the univariate tests on each of the four satisfaction items revealed no significant effects either.

Questionnaire concerning teachers' perception of workload: As mentioned above, 3 subscales were distinguished with regard to teachers' perception of workload as a result of implementing the innovative instructional approach. The overall mean scores in Table 4 reveal that teachers experience little workload in reference to settling in the innovative approach and preparing the students. Little to moderate workload was experienced with regard to the implementation and organization of peer tutoring activities. In addition,

teachers perceived the peer tutoring activities as a moderate to strong vehicle to build in differentiation according to pace or content for reading dyads of different abilities. A 2×2 MANOVA using the three subscales as the dependent measures, indicated no significant differences between the experiences of second- and fifth-grade teachers on the one hand ($F(3, 11) = 1.03$; $p = 0.416$) and the intensively coached teachers and teachers attending the restricted in-service professional development on the other hand ($F(3, 11) = 2.78$; $p = 0.091$). No significant interaction effect of grade level and condition was found either ($F(3, 11) = 1.45$; $p = 0.115$). As indicated in Table 5, the univariate tests, however, do reveal a significant condition effect on teachers' perceived workload in reference to settling in the innovative approach and student preparation, as well as a significant interaction effect between grade level and condition on the

Table 5

F-values and effect sizes of the Univariate Tests in the 2×2 MANOVA's performed on the teacher questionnaire variables

Questionnaire variables	Factors		
	Grade level (df = 1)	Condition (df = 1)	Grade level \times Condition (df = 1)
<i>Satisfaction</i>			
Overall	1.01 (0.06)	1.29 (0.07)	1.29 (0.07)
Information/counselling	0.00 (0.00)	0.07 (0.00)	0.39 (0.02)
Manual	0.14 (0.01)	0.02 (0.00)	0.02 (0.00)
Student materials	1.74 (0.10)	0.38 (0.02)	4.08 (0.20)
<i>Workload</i>			
Settle in/preparation	0.90 (0.06)	6.22* (0.32)	0.74 (0.05)
Organization tutoring	3.34 (0.21)	1.49 (0.10)	0.01 (0.00)
Chance to differentiate	0.00 (0.00)	2.17 (0.14)	4.88* (0.27)
<i>Student progress</i>			
Reading comprehension	0.45 (0.04)	2.47 (0.017)	7.27* (0.38)
Strategy knowledge	20.90** (0.64)	9.29* (0.44)	9.29* (0.44)
Strategy use	0.00 (0.00)	4.18 (0.26)	2.94 (0.20)
Spontaneous strategy use	0.20 (0.02)	0.20 (0.02)	0.45 (0.04)
General social skills	0.05 (0.00)	1.07 (0.09)	0.19 (0.02)
Social skills in tutoring	3.19 (0.23)	0.17 (0.02)	1.53 (0.12)
Social relationships	0.15 (0.01)	0.03 (0.00)	0.25 (0.02)
General task approach	0.51 (0.04)	4.24 (0.26)	0.20 (0.02)
Tutoring tasks approach	0.97 (0.07)	0.00 (0.00)	0.18 (0.02)

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.Note: Values in parentheses are effect sizes (partial η^2).

impression of opportunities to build in differentiation according to pace or content. The partial η^2 as measure of effect size more specifically indicates that, respectively, 32% and 27% of the variance in the dependent variable can be attributed to the experimental condition and the interaction between grade level and condition. In particular, it appears that teachers attending the restricted in-service development course experience a higher workload with regard to the preparation within the framework of the innovative approach. As to the interaction effect, fifth-grade teachers attending the restricted in-service development consider the organization of the peer tutoring activities more strongly as an effective opportunity to build in and increase differentiation according to pace or content for children of different abilities.

Questionnaire concerning teachers' appraisal of student progress: In reference to teachers' appraisal of student progress as a result of implementing the innovative instructional approach, different com-

ponents were distinguished, respectively, assessing students' progress in the field of reading comprehension and fluency, social and interaction skills, approach to assignments, and attitude towards reading. As regards the cognitive student outcomes, the overall mean scores presented in Table 4 reveal that teachers report a moderate to strong improvement of students' reading comprehension and fluency skills. A 2×2 MANOVA was performed on the four reading comprehension items. The multivariate tests showed a significant grade level ($F(4, 9) = 5.99$; $p = 0.012$), condition ($F(4, 9) = 6.29$; $p = 0.011$), and interaction effect ($F(4, 9) = 4.83$; $p = 0.023$), generally revealing that teachers in the restricted in-service professional development condition, especially those teaching fifth grade, report larger student advancement in the field of reading comprehension skills than the intensively coached teachers. As Table 5 indicates, some univariate tests reveal significant effects as well. More specifically, fifth-grade teachers and

teachers in the restricted in-service development condition report a significantly larger student progress in the field of knowledge of reading comprehension strategies than second-grade teachers and intensively coached teachers (grade main effect: partial $\eta^2 = 0.64$; condition main effect: partial $\eta^2 = 0.44$; grade \times condition interaction effect: partial $\eta^2 = 0.44$). In addition, fifth-grade teachers attending the restricted in-service development report a larger increase in general reading comprehension skills than their intensively coached counterparts. The partial η^2 more specifically indicates that 38% of the variance in teachers' reports of students' progress in reading comprehension can be attributed to the interaction between grade level and condition.

To investigate potential grade level or condition differences with regard to teachers' appraisal of students' reading fluency progress, a two way ANOVA was performed on the reading fluency item. The analysis indicated no significant grade level ($F = 0.62$; $df = 1$; $p = 0.442$) or condition main effect ($F = 1.83$; $df = 1$; $p = 0.195$), nor was there a significant interaction effect ($F = 2.14$; $df = 1$; $p = 0.162$).

With regard to teachers' perception of progress in the field of students' social and interaction skills, three subscales were distinguished. The overall mean scores presented in Table 4 indicate that teachers report a moderate to strong positive evolution in students' social relationships and friendships, as well as a moderate to strong improvement of their social and interaction skills, as a result of implementing peer tutoring activities within the framework of the innovative approach to reading comprehension instruction. The multivariate tests of a 2×2 MANOVA on the three subscales revealed no significant grade level ($F(3,9) = 1.84$; $p = 0.210$), condition ($F(3,9) = 2.21$; $p = 0.156$), nor interaction effect ($F(3,9) = 1.51$; $p = 0.278$). As Table 5 reveals, univariate tests revealed no significant differences either.

In reference to teachers' perception of progress in students' approach to assignments, two distinct subscales were developed. The overall mean scores presented in Table 4 indicate that teachers generally report a moderate to strong positive evolution in students' responsibility, self-confi-

dence, and autonomy when approaching peer tutoring assignments, as well as other tasks in general. A 2×2 MANOVA on the two subscales was performed. The multivariate tests revealed no significant grade level ($F(2,11) = 1.46$; $p = 0.275$), condition ($F(2,11) = 2.83$; $p = 0.102$), nor interaction effect ($F(2,11) = 0.38$; $p = 0.695$). As indicated in Table 5 the univariate tests on each of the three subscales revealed no significant effects either.

Finally, to investigate grade level or condition effects on teachers' appraisal of students' progress with regard to attitude towards reading, a two-way ANOVA was performed. The analysis indicated no significant grade level ($F = 0.38$; $df = 1$; $p = 0.546$) or condition main effect ($F = 1.38$; $df = 1$; $p = 0.255$), nor was there a significant interaction effect ($F = 0.38$; $df = 1$; $p = 0.546$).

5.2. Student outcomes

The student outcome data have a clear hierarchical structure with students nested within a smaller number of classes. More specifically, a repeated measures design was adopted, in which the three measurement occasions were considered as a distinct level within the students. Consequently, a three-level hierarchical structure arises: the three measurement occasions (level 1) are clustered within students (level 2), in their turn nested within classes (level 3). Within this framework student data were analyzed using hierarchical linear modeling, for these models are specifically geared to take statistical account of data with clustered structure (Goldstein, 1995). For each dependent variable, a stepwise procedure was followed to build appropriate models and to test the research hypothesis. Separate analyses were done for second and fifth graders. The parameters of the models were estimated using the restrictive iterative generalized least squares (RIGLS) estimation procedure of the software MLwiN (Rasbash et al., 1999). Since parsimonious models are preferred, only significant estimates ameliorating the model were retained.

Reading comprehension achievement: Besides the effects for the measurement occasions, the final repeated measures model with regard to second

and fifth graders' reading comprehension achievement included explanatory variables such as gender, mother tongue and the number of years students are behind at school. For second grade reading fluency scores were also included in the model. More specifically, it appears that second-grade boys perform significantly lower than girls with a fixed arrear at each measurement occasion. A similar effect was found for being a non-native Dutch speaker in fifth grade. On the average second-grade non-native speakers also perform significantly lower at pre-test, but they appeared to catch up significantly in the course of second and third grade. Further, the number of years second and fifth graders are behind at school has been found inversely proportional to their comprehension scores, with a fixed impact for all the measurement occasions. Finally, as regards the impact of reading fluency it can be concluded that second graders with higher pre-test fluency scores on the average have higher pre-test reading comprehension scores as well. Nevertheless, the relationship between reading fluency and reading comprehension becomes weaker at the post- and retention test. The class-level total amount of time spent on reading comprehension instruction and class size were also introduced as potential predictors, but did not reveal significant effects in none of the analyses.

Having controlled for the aforementioned significant explanatory variables, the results revealed that at pre-test second- ($\chi^2 = 0.17$, $df = 1$, $p = 0.684$) and fifth-grade students ($\chi^2 = 3.75$, $df = 1$, $p = 0.053$) with teachers in the restricted professional development condition did not significantly differ from their peers with teachers in the intensive coaching condition. Moreover, no significant differences were found between both research conditions with regard to students' progress in reading comprehension achievement from pre- to post-test in second ($\chi^2 = 0.83$, $df = 1$, $p = 0.362$) and fifth grade ($\chi^2 = 0.06$, $df = 1$, $p = 0.805$), nor from pre-test to retention test in second ($\chi^2 = 1.51$, $df = 1$, $p = 0.219$) and fifth grade ($\chi^2 = 0.09$, $df = 1$, $p = 0.764$).

Reading fluency achievement: Second graders were also tested with regard to their reading fluency achievement. The final multilevel model

indicated that second-grade boys performed significantly lower than girls with a fixed arrear at each measurement occasion and that at pre-test no significant differences arose between students with teachers in the restricted course or intensive coaching condition ($\chi^2 = 0.17$, $df = 1$, $p = 0.685$). In addition, no differential reading fluency progress was found from pre-test to post- ($\chi^2 = 1.06$, $df = 1$, $p = 0.303$) and retention test ($\chi^2 = 0.17$, $df = 1$, $p = 0.683$).

Use of reading comprehension strategies: With regard to students' report of strategy use, similar models were built up. It appeared that at each measurement occasion second-grade boys reported significantly less strategic reading than girls. No significant pre-test differences between the research conditions were found, neither for second ($\chi^2 = 0.24$, $df = 1$, $p = 0.622$), nor for fifth grade ($\chi^2 = 3.29$, $df = 1$, $p = 0.070$). Moreover, no differential increase or decrease in the reports of reading strategy use of students from both conditions was found from pre- to posttest in second ($\chi^2 = 2.55$, $df = 1$, $p = 0.110$) and fifth grade ($\chi^2 = 2.626$, $df = 1$, $p = 0.105$), nor from pre-to retention test in second ($\chi^2 = 0.00$, $df = 1$, $p = 0.964$) and fifth grade ($\chi^2 = 3.53$, $df = 1$, $p = 0.060$).

Self-efficacy perceptions and related causal attributions: The final multilevel models with regard to students' self-efficacy measures reveal that besides the dummies representing the measurement occasions second graders' gender and fifth graders' mother tongue yielded significant estimates and improvements of the models. More specifically, at each measurement occasion second-grade boys reported significantly less thoughts with regard to success attributions and positive self-efficacy perceptions than girls. As concerns the fifth graders, at each measurement occasion non-native students reported significantly more thoughts with regard to failure attributions and negative self-efficacy perceptions than Dutch-speaking students.

Having controlled for the aforementioned explanatory variables, the models with regard to second and fifth graders' thoughts relating to success attributions and positive self-efficacy perceptions towards reading indicate no significant pre-test differences between both research

conditions, neither for second ($\chi^2 = 0.00$, $df = 1$, $p = 0.964$), nor for fifth grade ($\chi^2 = 0.69$, $df = 1$, $p = 0.407$). Moreover, no significantly differential evolution was found from pre- to post-test in second ($\chi^2 = 0.16$, $df = 1$, $p = 0.686$) and fifth grade ($\chi^2 = 0.09$, $df = 1$, $p = 0.762$), nor from pre- to retention test in second ($\chi^2 = 1.54$, $df = 1$, $p = 0.214$) and fifth grade ($\chi^2 = 0.82$, $df = 1$, $p = 0.365$). The models with regard to second and fifth graders' thoughts relating to failure attributions and negative self-efficacy perceptions indicate similar results: no significant pre-test differences between the second- ($\chi^2 = 3.30$, $df = 1$, $p = 0.069$) and fifth-grade ($\chi^2 = 1.27$, $df = 1$, $p = 0.260$) research conditions; and no significantly differential evolution from pre- to post-test in second ($\chi^2 = 0.00$, $df = 1$, $p = 0.950$) and fifth grade ($\chi^2 = 2.83$, $df = 1$, $p = 0.093$), nor from pre- to retention test in second ($\chi^2 = 1.32$, $df = 1$, $p = 0.250$) and fifth grade ($\chi^2 = 3.48$, $df = 1$, $p = 0.062$).

6. Discussion and conclusion

The purpose of the present study was to compare the effectiveness of two variants of professional development of teachers, developed to gear reading comprehension instruction in elementary schools to the findings and research-based successful practices emanated from empirical intervention research. The variants were characterized by different amounts of support and assistance lent to the teachers. More specifically, the efficacy of a year-round elaborated coaching of the teachers (35 h) spread over the entire school year was equated to a more restricted 13-h in-service course by using a questionnaire assessing teachers' experiences with the professional development course and with the implementation of the innovations, as well as student outcomes as a result of the renewed instructional practice.

With regard to student outcomes the results of the multilevel linear regression analyses can be summarized easily: both conditions for professional development of teachers were equally effective in changing students' reading comprehen-

sion, reading fluency, use of reading strategies, and self-efficacy perceptions towards reading from pre- to post- and retention test. These results confirm the hypothesis that student outcomes are comparable in both the intensive year-round coaching and the restricted in-service teacher development course.

As to the experiences reported by the teachers, the analyses generally confirmed the central research hypothesis. Teachers in both conditions merely reported positive experiences. A picture emerges that both second- and fifth-grade teachers were strongly satisfied with their participation in the development course, with the offered information and counselling, and with the provided manual and student materials. Moreover, their reports indicate that implementing the innovative reading comprehension instruction as a result of attending one of the teacher development courses demanded only little to moderate additional workload. Teachers also perceived the organization of peer tutoring activities as a moderate to strong vehicle to build in differentiation according to pace or content for reading dyads of different abilities. As an additional advantage they report moderate to strong student improvement in the field of reading comprehension and fluency skills, social and interaction skills, social relationships and friendships, attitude towards reading, responsibility, self-confidence, and autonomy when approaching assignments as a result of the implementation of the innovative reading comprehension instruction.

With regard to the differences in the experiences reported by the teachers, the results revealed that second- and fifth-grade teachers in both conditions were equally satisfied with the course they attended and with the provided manual and student materials. Except for teachers' assessment of students' reading comprehension progress, no differences were detected either with regard to teachers' appraisal of students' progress. As concerns the difference in the evaluation of students' reading comprehension progress, the results more specifically favored the restricted in-service professional development condition, for teachers in this condition, especially those teaching fifth grade, reported larger student improvement

than the intensively coached teachers. This result contrasts with the finding that with respect to the actual student learning gains no differences were found between the two strategies for professional development of teachers. Presumably, the finding that teachers in the more restricted in-service development condition perceived higher learning progress in their students' reading comprehension can be attributed to the newness of the peer tutoring activities for those teachers, whereas teachers in the intensive coaching condition implemented the entire innovative approach (i.e. explicit reading strategies instruction and peer tutoring activities) for the second consecutive year, which could have led to a habituation effect. For teachers in the more restricted in-service development the implementation of peer tutoring activities in their classrooms was something completely new since the previous year they had only implemented part of the innovation, more particularly the explicit teaching of reading strategies. It is conceivable that the teachers' enthusiasm for this new experience led them to perceive greater student progress.

Finally, the analyses with regard to the perceived workload as a result of implementing the innovative instructional approach did reveal significant differences between both teacher development conditions. More specifically, it appeared that teachers attending the restricted in-service development course experienced a higher workload with regard to settling in the innovative approach and preparing the students for practicing the application of reading strategies in peer tutoring dyads. Apparently, this finding goes together with the fact that in the restricted in-service teacher development only three 3-h preparatory meetings with the teachers were organized, while in the intensive coaching condition 8 local meetings took place. So, the former group of teachers needed more additional time apart from the in-service sessions to study and prepare for the intended innovative instructional approach. It is conceivable that individual extra preparation time is perceived as causing more extra workload compared to the comfort of attending coaching sessions in which everything can be discussed with the external coach. But also the fact that for the

teachers in the more restricted in-service condition implementing peer tutoring was something new—while teachers in the intensive coaching condition had previous experience in it—could be responsible for this higher perception of workload.

Notwithstanding the higher report of pressure of preparatory work, fifth-grade teachers attending the restricted in-service professional development considered the organization of the peer tutoring activities more strongly as an effective opportunity to build in and increase differentiation according to pace or content for children of different abilities. On the basis of the content and structuring of both strategies for professional development of teachers, however, no evident explanation of this difference can be formulated. But again, the fact that compared to the teachers in the other condition, implementing peer tutoring was something new for teachers in the restricted in-service condition may have caused them to appreciate in greater extent some of the pedagogical advantages of this new teaching method.

In conclusion, the results concerning the experiences reported by the teachers generally confirm the research hypothesis that teachers' experiences in the restricted in-service course are comparable to those in the intensive coaching condition. For only three elements in the teacher questionnaire significant differences could be determined, of which two differences favored the restricted in-service course. The third difference suggests that teachers in the restricted in-service course need to spend more time and energy on settling in the innovative approach and preparing the students. The additional workload, as perceived by teachers in the restricted in-service development does; however, not undermine teachers' satisfaction with attending the development course. With regard to the affirmation of the research hypothesis on teachers' reported experiences, it has to be taken into account that the absence of important significant differences between the restricted and intensive coaching condition might be due to the rather small sample size of 30 participating teachers, which was the unit of analysis in the analyses of variances. In this respect it might be important to reply the study with a larger number

of participating teachers in the different variants of teachers' professional development.

Finally, some limitations of the present study ought to be noted. First, it should be mentioned that teachers were asked to deliver or mail the questionnaire with regard to their experiences with the program to the course instructor. Therefore, it is possible that socially desirable responses were elicited from some of the respondents. In this respect it would be interesting to document the validity of the questionnaire and assess teachers' tendency to social desirability in answering the questionnaire, for example by using in-depth interviews as an additional method of data collection. Unfortunately, this was practically impossible within the scope of the present study.

A second restriction is connected to the fact that the first author of the article who conducted both teacher development programs was already experienced in intensively coaching teachers to innovate their reading comprehension instruction from former research (Van Keer, 2004; Van Keer & Verhaeghe, 2003). In this respect, the objection can be raised that because of the previous elaborated coaching and counselling experience, the quality of the restricted in-service teacher development was higher than can be expected in the case of a novice or less experienced counsellor of the teacher development sessions. Further research is necessary to investigate whether the quality of the present 13-h in-service teacher program and the connected teacher satisfaction can be equalled by disseminating the innovative approach on a broad scale through regular in-service professional development authorities. Since these authorities often prefer to assemble participants from different school teams, additional research should in this respect also try to verify whether the organization of the three 3-h preparatory local school meetings can be replaced by collective meetings without losing quality.

A third comment can be made on the fact that only experiences reported by the teachers and student outcome measures were used to compare the effectiveness of both teacher development programs. It can be argued that it is as important to gather information about the extent to which the implementation of the innovations has been

carried out correctly and in accordance with the intention of the teacher development program. Notwithstanding the fact that regular discussions with the teachers took place and observations in the classes were executed, a comment can be made on the fact that within the scope of this study, it was practically impossible to collect quality of the implementation and treatment fidelity data systematically. Therefore, subsequent research should complete the data gathered in the present study with evaluations of the treatment integrity of the innovative reading comprehension instruction approach as a result of the different teacher development programs. In this respect, the use of a structured observations by a double-blind observer, unacquainted with the experimental condition teachers are assigned to, could be useful. This kind of detailed and rather qualitatively in-depth analysis will be a supplementary source of information to shed light on the shifts in teachers' practice and the quality of the different professional development programs and on teachers' implementation of the educational innovation in pursuance of the program.

Although more research is needed and notwithstanding the restrictions, the outcomes of the present study clearly corroborated that a compact in-service teacher development aimed at changing instructional content and teacher behaviors in the classrooms was equally effective as the year-round coaching of teachers. Taken into account previous studies referred to above, the similar effectiveness of the restricted in-service teacher development as compared to the intensive year-round coaching can be attributed to the combination of different components, notably the focused presentation of information about the background of the innovative instructional approach, the identification and discussion of the intended instructional practice, the feasible and elaborated manual with lesson scenarios and all necessary teacher and student materials, the demonstration of relevant instructional behaviors to be implemented by means of accurately selected video fragments, the structured opportunities for reflection, and follow-up sessions with in-class supervised practice, on-site consultation, and performance feedback to the teachers. It cannot be said from this study that one aspect of

the restricted professional development was more powerful than the others, or that one aspect preceded another. In the present study they combined to have an impact on teachers' experiences and student outcomes that was equally effective as a more intensive year-round coaching of teachers.

References

- Aarnoutse, C.A.J. (1995). *Probleemidentificatie en aanzet voor een actieplan taal* [Problem identification and initiative towards an action plan for language]. Nijmegen, The Netherlands: Vakgroep Onderwijskunde.
- Aarnoutse, C. A. J., & Van Leeuwe, J. F. J. (2000). Development of poor and better readers during the primary school. *Educational Research and Evaluation*, 6, 251–278.
- Aarnoutse, C. A. J., & Weterings, A. C. E. M. (1995). Onderwijs in begrijpend lezen [Education in reading comprehension]. *Pedagogische Studiën*, 72, 82–101.
- Almasi, J. F. (1996). The nature of fourth graders' socio-cognitive conflicts in peer-led and teacher-led discussion of literature. *Reading Research Quarterly*, 30, 314–351.
- Alvermann, D. E. (2000). Classroom talk about texts: Is it dear, cheap or a bargain at any price? In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning. Fostering comprehension in the middle grades* (pp. 170–192). New York: Teachers College Press.
- Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353–394). New York: Longman.
- Ball, D. L., & Cohen, D. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, 25, 6–8.
- Bell, L. (1991). Approaches to the professional development of teachers. In L. Bell, & C. Day (Eds.), *Managing the professional development of teachers* (pp. 3–22). Buckingham, United Kingdom: Open University Press.
- Bentz, J. L., & Fuchs, L. S. (1996). Improving peers' helping behavior to students with learning disabilities during mathematics peer tutoring. *Learning Disability Quarterly*, 19, 202–215.
- Bradley, H., Conner, C., & Southworth, J. (Eds.). (1994). *Developing teachers, developing schools*. London: David Fulton.
- Brand-Gruwel, S., Aarnoutse, C. A. J., & Van Den Bos, K. P. (1998). Improving text comprehension strategies in reading and listening settings. *Learning and Instruction*, 8, 63–81.
- Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology*, 88, 18–37.
- Brus, B.Th. (1969). *Eén-minuut-test* [One minute test]. Nijmegen, The Netherlands: Berkhout.
- Burbank, M. D., & Kauchak, D. (2003). An alternative model for professional development: Investigations into effective collaboration. *Teaching and Teacher Education*, 19, 499–514.
- Butler, D. L., Lauscher, H. N., Jarvis-Selinger, S., & Beckingham, B. (2004). Collaboration and self-regulation in teachers' professional development. *Teaching and Teacher Education*, 20, 435–455.
- Clarke, D. J., & Hollingsworth, H. (2000). Seeing is understanding: Examining the merits of video and narrative cases. *Journal of Staff Development*, 21, 40–43.
- Clarke, D. J., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18, 947–967.
- De Corte, E., Verschaffel, L., & Van De Ven, A. (2001). Improving text comprehension strategies in upper primary school children: A design experiment. *British Journal of Educational Psychology*, 71, 531–559.
- Djalil, A., & Anderson, L. W. (1989). The impact of a research-based teacher training program on Indonesian teachers, classrooms and students. *Teaching and Teacher Education*, 5, 165–178.
- Dole, J. A. (2000). Explicit and implicit instruction in comprehension. In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning. Fostering comprehension in the middle grades* (pp. 52–69). New York: Teachers College Press.
- Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research*, 61, 239–264.
- Englert, C. S., & Tarrant, K. L. (1995). Creating collaborative cultures for educational change. *Remedial and Special Education*, 16, 325–336.
- Fuchs, L. S., & Fuchs, D. (2000). Building student capacity to work productively during peer-assisted reading activities. In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning. Fostering comprehension in the middle grades* (pp. 95–115). New York: Teachers College Press.
- Fuchs, L. S., Fuchs, D., Bentz, J., Phillips, N. B., & Hamlett, C. L. (1994). The nature of students interactions during peer tutoring with and without prior training and experience. *American Educational Research Journal*, 31, 75–103.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., Karns, K., & Dutka, S. (1997). Enhancing students' helping behavior during peer-mediated instruction with conceptual mathematical explanations. *The Elementary School Journal*, 97, 223–249.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206.
- Fukkink, R., Van der Linden, S., Vosse, A., Vaessen, K. (1997). *Stap door! Handleiding* [Step out! Manual]. Utrecht, The Netherlands: Sardes.

- Fullan, M., & Stiegelbauer, S. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Gersten, R., Vaughn, S., Deshler, D., & Schiller, E. (1997). What we know about using research findings: Implications for improving special education practice. *Journal of Learning Disabilities*, 30, 466–476.
- Goldstein, H. (1995). *Multilevel statistical models*. London: Edward Arnold.
- Harchik, A. E., Sherman, J. A., Sheldon, J. B., & Strouse, M. C. (1992). Ongoing consultation as a method of improving performance of staff members in a group home. *Journal of Applied Behavior Analysis*, 25, 599–610.
- Hartman, H. J. (2001). Developing students' metacognitive knowledge and skills. In H. J. Hartman (Ed.), *Metacognition in learning and instruction* (pp. 33–68). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Hawley, W., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling-Hammond, & G. Sykes (Eds.), *Teaching as a learning profession: Handbook of policy and practice* (pp. 127–150). San Francisco: Jossey-Bass.
- Joyce, B., & Showers, B. (1980). Improving in-service training: The message of research. *Educational Leadership*, 37, 379–385.
- Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational Leadership*, 40, 4–10.
- Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *The Elementary School Journal*, 96, 275–293.
- Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth-grade classrooms. *The Elementary School Journal*, 99, 3–21.
- Leach, D. J., & Ingram, K. L. (1989). The effects of information and feedback on teachers' classroom behavior and students' academic engaged time. *Educational Psychology*, 9, 167–184.
- Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, 15, 129–151.
- Loucks-Horsley, S., Hewson, P., Love, N., & Stiles, K. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.
- Maes, F., Vandenbergh, R., Ghesquière, P., & Luyckx, E. (1997). *Het diagnostiserend en remediërend vermogen van lagere scholen. Eindraport*. [Elementary schools' capability in diagnosing and remediating. Final Research Report] Brussels, Belgium: Ministry of the Flemish Community, Department of Education.
- Mathes, P. G., & Fuchs, L. S. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review*, 23, 59–80.
- McCutchen, D., & Berninger, V. W. (1999). Those who know, teach well: Helping teachers master literacy-related subject-matter knowledge. *Learning Disabilities Research and Practice*, 14, 215–226.
- Miller, E. (1998). The old model of staff development survives in a world where everything else has changed. In R. Tovey (Ed.), *Professional development, Harvard education letter focus series no. 4* (pp. 1–3). Cambridge, MA: Harvard Education Letter.
- Norman, A., Sprinthall, A. J., & Ties-Sprinthall, L. (1996). Teacher professional development. In J. Sikula (Ed.), *Second handbook of research on teacher education* (pp. 666–703). New York: Macmillan.
- Owens, L., & Barnes, J. (1992). *Learning preference scales: Handbook and test master set: Teachers, students, parents*. ACER: Hawthorn, Victoria.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117–175.
- Palincsar, A. S., Magnusson, S., Marano, N., Ford, D., & Brown, N. (1998). Designing a community of practice: Principles and practices of the GISML community. *Teaching and Teacher Education*, 14, 5–19.
- Paris, S. G., & Oka, E. R. (1986). Children's reading strategies, metacognition and motivation. *Developmental Review*, 6, 25–56.
- Perry, N. E., Walton, C., & Calder, K. (1999). Teachers developing assessments of early literacy: A community of practice project. *Teacher Education and Special education*, 22, 218–233.
- Pressley, M. (2000). Comprehension instruction in elementary school: A quarter-century of research progress. In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning. Fostering comprehension in the middle grades* (pp. 32–51). New York: Teachers College Press.
- Pressley, M., & Allington, R. (1999). What should reading instructional research be the research of? *Issues in Education*, 5, 1–35.
- Pressley, M., El-Dinary, P. B., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J., & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal*, 92, 513–555.
- Pressley, M., Wharton-McDonald, R., Hampson, J. M., & Echevarria, M. (1998). The nature of literacy instruction in ten grade-4/5 classrooms in upstate New York. *Scientific Studies of Reading*, 2, 159–191.
- Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., & Draper, D. (1999). *A user's guide to MLwiN*. London: Institute of Education.
- Reezigt, G. J. (1993). *Effecten van differentiatie op de basisschool*. [Effects of differentiation in elementary schools]. Groningen, The Netherlands: RION.
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of research. *Review of Educational Research*, 64, 479–530.
- Sandholtz, J. H. (2002). Inservice training or professional development: Contrasting opportunities in a school/university partnership. *Teaching and Teacher Education*, 18, 815–830.

- Showers, B. (1990). Aiming for superior classroom instruction for all children: A comprehensive staff development model. *Remedial and Special Education*, 11, 35–39.
- Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P. G., & Hodge, J. P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *The Elementary School Journal*, 95, 387–408.
- Staphorsius, G., & Krom, R. (1996). *Toetsen Begrijpend Lezen* [Reading Comprehension Tests]. Arnhem, The Netherlands: Cito.
- Stein, M. K., Schwan Smith, M., & Silver, E. A. (1999). The development of professional developers: Learning to assist teachers in new settings in new ways. *Harvard Educational Review*, 69, 237–269.
- Van Den Broek, P., & Kremer, K. E. (2000). The mind in action: What it means to comprehend during reading. In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning. Fostering comprehension in the middle grades* (pp. 1–31). New York: Teachers College Press.
- Van Keer, H. (2002). *Een boek voor twee. Verwerving van strategieën voor begrijpend lezen via peer tutoring*. Handleiding. [A book for two. Acquisition of reading comprehension strategies through peer tutoring. Manual] Antwerpen, Belgium: Garant.
- Van Keer, H. (2004). Fostering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring. *British Journal of Educational Psychology*, 74, 37–70.
- Van Keer, H., & Verhaeghe, J. P. (2003). Effects of explicit reading strategies instruction and peer tutoring in second and fifth graders. In E. De Corte, L. Verschaffel, N. Entwistle, & J. Van Merriënboer (Eds.), *Powerful learning environments: Unravelling basic components and dimensions* (pp. 211–228). Oxford: Elsevier.
- Van Keer, H., & Verhaeghe, J. P. (2004). *Effects of peer tutoring on second graders' reading fluency development*. Manuscript submitted for publication.
- Veenman, S., Van Tulder, M., & Voeten, M. (1994). The impact of in-service training on teacher behavior. *Teaching and Teacher Education*, 10, 303–317.
- Verhoeven, L. (2003). *Lezen met begrip 1* [Reading with comprehension 1]. Arnhem, The Netherlands: Cito.
- Wade, R. K. (1985). What makes a difference in in-service teacher education? A meta-analysis of research. *Educational Leadership*, 42, 48–54.
- Wheldall, K., Merrett, F., & Borg, M. (1985). The behavioural approach to teaching package (BATPACK): An experimental evaluation. *British Journal of Educational Psychology*, 55, 65–75.